International Journal of Artificial Intelligence
p-ISSN: 2407-7275, e-ISSN: 2686-3251

Original Research Paper

Automatic Pose Recognition in Basketball Videos Using Entropy, Mean and Standard Deviation

Aliga Paul Aliga[1*], Joshua Ojo Nehinbe[2], Kingsley Eghonghon Ukhurebor[3]

[1] *Department of Computer Science, Ambrose Alli University. Edo State, Nigeria.*
[2] *Department of Computer Science, Edo State University. Edo State, Nigeria.*
[3] *Department of Physics, Edo State University. Edo State, Nigeria.*

Article History
Received:
30.08.2025

Revised:
15.09.2025

Accepted:
28.09.2025

***Corresponding Author:**
Aliga Paul Aliga
Email
aliga.paul@aauekpoma.edu.ng

Abstract: Most existing models for automatic action recognition in basketball videos lack privacy-friendly analytics, versatility and explainability. So, coaches, players and analysts often invest substantial resources by relying heavily on visual appearance, ball tracking and court context. Unfortunately, this method can be resource-intensive and potentially susceptible to unforeseeable intrusions. This study proposes an entropy-based analytical model for automatic recognition of key basketball actions, designed to optimize the video review process to address the above limitations. The model is implemented with Python programming language to analyze entropy arrays, the mean and standard deviation values derived from 22 basketball game videos. Evaluation suggests that the model flagged basketball_Video2, Video3 and Video9 as containing key moments deserving closer inspection. This has successfully reduced the input datasets to just three critical videos (with mean and standard deviation pairs of 1.96 & 0.33, 2.05 & 0.31, and 1.94 & 0.20) that warrant detailed examination. This targeted filtering significantly improves review efficiency by conserving time and resources and effectively eliminated 19 videos deemed redundant or of lower priority. The approach demonstrates high precision in identifying impactful gameplay moments and addresses a long-standing challenge with workload reduction in basketball analytics without sacrificing review accuracy. Consequently, this method not only supports privacy-conscious analytics but also provides coaches, players and sports analysts with a more focused, resource-efficient framework they can adopt for performance evaluation and strategic decision-making in basketball.

Keywords: Action Recognition, Basketball Videos, Entropy, Pose Extraction from Basketball Videos, Privacy-Friendly Analytics.

1. Introduction

Manual extraction of pose within the frames of basketball videos presents several challenges [1]. This issue is recently generating serious concerns with the dynamic nature of posture transitions and immense benefits of human pose analysis in the recognition, understanding and usage of various actions that are within basketball video footage after gameplay [2]. The general consensus is that posture transitions should be algorithmically scored to indicate the importance and relevance of video segments for detailed analysis with aim to reduce workload without compromising the depth of player behavior assessment and the complexity of movement in basketball videos [3] [4]. Consequently, automatic extraction of human pose from frames of basketball videos has constituted an active area of research since the last two decades [5] [6] [7]. There are specific problems that these ideas can address especially if the model is accurately designed to analyze human pose in basketball videos.

Experience shows that entropy arrays can play significant role in action recognition using pose dynamics in basketball videos [8]. The argument is that entropy arrays can distinguish between low-entropy actions like standing and walking against high-entropy actions like jumps and rapid cuts where players must vigorously dissipate energy and efforts to achieve best performance or regarded as skillful players. Player activity and intensity estimation vary across basketball videos but it is practically inefficient to extract them with the use of manual methods. Entropy is often associated with variability or uniformity of actions. High entropy often reflects dynamic postures and complex physical activities. These facts make entropy a valuable indicator for coaching and injury prevention. Failure to recognize such patterns may lead to recurring injuries and missed coaching opportunities which can severely hinder team performance. Therefore, it is imperative to detect key events such as shooting, passing, etc. and establish the relationship between them and temporal spikes in pose entropy among basketball events. Coaches and video analysts often face heavy workloads and challenges requiring trade-offs when reviewing large volumes of footage to assess player styles and skill levels among basketball players. Besides, comparing entropy patterns to distinguish amateur from professional play has remained a complex and resource-intensive challenge over the years. Many existing basketball video analysis tools are proprietary, expensive and often lacking advanced features such as entropy-based pose segmentation. Their reliance on manual fine-tuning to classify video segments [9] into gameplay phases such as idle, transition and defense remains a significant limitation hindering their scalability and broader adoption as effective context-aware analysis tools.

Moreover, one of the core objectives of the model is to implement a model that can use MediaPipe and entropy to automatically extract human pose landmarks from each frame in basketball videos to lessen the above challenges. Secondly, the model should be able to automate temporal segmentation and feature extraction by dividing each of its input basketball videos into 1-second clips and extract pose landmarks over time that are present in them. Thirdly, the model intends to compute entropy of spatial (x, y) distributions of pose landmarks per clip and represent the level of variability, movement complexity and motion patterns of players in all the input basketball videos for human analysis, classification and visualization. Furthermore, the above objectives are implemented with Python programming language, 22 basketball videos, entropy arrays and Python's associated libraries such as os, numpy, cv2, pathlib, mediapipe and sub- package like *scipy.stats*.

Moreover, one of the substantial contributions of this paper is its ability to support and form the basis of research in action recognition, player behavior analysis and movement complexity studies in basketball. The paper specifically addresses action recognition in basketball videos by using multi-metric arrays involving entropy, mean and standard deviation to categorize and distinguish pose dynamics. The paper demonstrates how player activities and intensity can be used to detect players of different skill levels and suggest how the findings reported in this paper can be used for effective coaching, style or skill differentiation and injury prevention of players. The paper pragmatically illustrates how to succinctly categorize and segment gameplay in numerous basketball videos into distinct review types, namely low, medium and high to suggest segments of heightened relevance and therefore warrant closer review.

2. Literature Review

2.1. Related Works

Automatic action recognition in basketball videos has garnered significant attention in recent years. There are numerous ways to categorize research that relate to pose recognition in basketball vides. Some authors leverage on deep learning, machine learning, neural networks, fusion techniques

International Journal of Artificial Intelligence, vol. 12, no. 2, pp. 101-115, December 2025. **DOI: 10.36079/lamintang.ijai-01202.908**

entropy characterization, motion analysis and sensor integration to classify actions in basketballs (e.g. player skill levels), enhancing performance evaluation and strategic planning of basketball games over the years. Some of scholars incorporate IMU classifier to gather human movements based on IMU sensor data [10]. IMU is a wearable device widely used in sports, healthcare, rehabilitation and biomechanics. The device uses Inertial Measurement Units (IMUs) to recognize, classify and provide feedback on human movements in a real-time manner.

1) Deep and Machine Learning-Based Action Recognition

Extraction of spatial, temporal and context with the aid of multi-stream deep learning has been proposed to explore pose estimation in basketball video analysis [11]. Though this the multi-stream architecture is comprehensive in understanding pose structure and dynamics, it does not deeply focus on segmentation. Also, the mode is likely to incur heavy computational load and resource usage during deployment. The model also lacks comparative baseline and it does not leverage on entropy profiles. Study that uses deep neural network model incorporated with a dynamic residual attention mechanism has been proposed to improve the accuracy of basketball action recognition [3]. Their method enhanced feature extraction and focused on key frames. The authors reported that their model achieved above 97% accuracy in recognizing basketball postures. Two-stream 3D Convolutional Neural Networks (CNNs) framework that [12] fuse global and local motion patterns with key visual information have been designed to perform semantic event recognition in basketball videos. Study shows that their approach demonstrated actions in basketball can be grouped into success and failure performance using NCAA dataset. Some authors integrated interactive system with Machine Learning (ML) including accuracy and recall scores to recognize and select the most suitable actions for basketball training systems [5].

2) Sensor-Based Action Recognition

Multiple sensors have been explored to optimize energy usage. The model used Multilayer [13] Parallel Long Short-Term Memory (MP-LSTM) for high-accuracy activity recognition and Deep Q-Network (DQN) for reinforcement learning. However, the model heavily relies on wearable sensor data. Apart from the fact that it does not incorporate video or visual pose information, it is limited to few motion-sensitive tasks. It also recorded lower accuracy due to noise from vigorous movement. The development of the EITNet has proved that IoT-enhanced framework can be used to explore real-time recognition of actions in basketball games [14]. They integrated EfficientDet for object detection with I3D for spatiotemporal feature extraction and TimeSformer for temporal analysis. It was reported that the model achieved a recognition accuracy of 92%. Sensor-driven model that utilized inertial measurement units (IMUs) in real-time manner has been proposed to detect basketball shooting motions [15]. This model was reported to have achieved high accuracy in classifying different types of shooting motions required to provide valuable insights into player performance.

3) Other Methods

Interactive system that collected acceleration and angular velocity of participants from wearable sensors [5] and used machine learning algorithms to classify their actions into passing, dribbling and shooting was another dimension to action recognition in basketball over the years. Studies proposed an ontology-based global and collective motion pattern algorithm for event classification in basketball videos [16]. This approach combined optical flow-based motion extraction with deep learning models to achieve a mean average precision of 58.10% on the NCAA+ dataset. The introduction of smart devices and Hang-Time HAR dataset have been proposed to benchmark models for evaluating basketball activity recognition [17]. The model used wrist-worn inertial sensors. Their input dataset includes 24 players across different countries. The model emphasizes variability of skill levels, cross-cultural and cross-skill-level analyses. Motion entropy was used as a feature [18] to characterize the complexity and variation of movement in sports videos and to support automatic segmentation of gameplay into meaningful events. However, their model focused mostly on sports video segmentation. Besides, the model was not customized for action recognition like dribbling, shooting and passing that are peculiar to basketball games. In essence, the above papers did not specifically employ entropy, mean and standard deviation as theoretical metrics for core feature extraction and decision logic in their approaches to action recognition in basketball videos.

Nonetheless, significant numbers of the above models rely heavily on appearance, court visuals and sensor data. Thus, they can potentially generate privacy concerns. Deep learning models such as 3D CNN and IMU classifiers are resource intensive. They often lack intuitive interpretation compare to entropy metrics such as mean and standard deviation that provide more transparent and interpretable signals regarding action recognitions in basketball videos. The above models were trained on narrow pose/action sets and controlled datasets without necessarily capturing the generic pose variability commonly found in modern basketball games. Above all, some of the existing models treat all clips equally without prioritization [19]. For these reasons, they might be ineffective when adapted to filter out redundant basketball videos with the view to reduce workload associated with critical review of numerous basketball videos without compromising accuracy for resource conservation.

2.2. Pose Extraction in Basketball Videos

Pose refers to a player's specific body position or posture during gameplay in the context of basketball videos. This can also indicate how a player is standing, moving or positioning himself (or herself) to perform actions within basketball game. Pose extraction in basketball videos refers to the process of using computer vision and techniques such as Machine Learning (ML) and Deep Learning (DL) to identify and analyze the precise body positions and movements of players throughout the game [20] [21] [22]. Pose extraction enables automatic detection of key body joints such as knees, elbows and shoulders across video frames. Understanding the pose dynamics can allow coaches, players and analysts to have deeper understanding of player actions, tactics and biomechanics recorded in basketball footage. Individual basketball players and teams habitually exhibit unique patterns of movement, behavior and body posture during gameplay. These patterns are largely influenced by many factors such as role within the team, playing style and skill level. Consequently, players can often be categorized based on these distinctive styles and their level of expertise. Nevertheless, how to effectively distinguish between the level of performance of amateur from professional performance particularly by critical analysis of movement and pose patterns has remained a complex and evolving challenge over the years for a number of reasons. One of the contributing factors to this complexity is the difficulty in comparing entropy with multi- metrics to explain pose patterns that signify the variability and unpredictability of poses and movements across different players. Another issue is the limited capability of that automate downstream tasks such as performance evaluation and tactical analysis tools to optimize coaching strategies.

Empirical reviews point out that modern coaching practices increasingly rely on post-game video reviews [23] [24]. For this reason, coaches and players must review basketball videos both of their team's individual performance and that of opponents to firmly develop strategies aimed at winning or at the very least, drawing competitive games. In this context, pose extraction from basketball videos has emerged as a valuable instrument for acquiring informative ideas to support decision making. We premise that by automatically identifying and isolating players' body positions across frames of basketball videos, analysts and Artificial Intelligence (AI) systems can split complex actions in basketball games into structured, simplified and easily interpretable segments.

These extracted pose clips have many potential applications such as performance analysis, coaching education, training optimization and the generation of automated game highlights rather watching redundant basketball videos [25] [26]. Redundant basketball videos denote video clips that contain very much similar and repetitive action sequences such that viewing one of them provides satisfactory information that can be suitably used for decision-making while tactically rendering the rest videos unnecessary. Watching these redundant clips can be time-consuming, inefficient and monotonous because they always fail to present (or convey) new and meaningful insights to the viewers. In performance analysis and coaching, minimizing such redundancy is crucial to maintaining engagement, effective time and resource management, improving efficiency and strongly focusing attention on unique and critical events (or moments) that truly inform strategy and development [27] [28]. On a whole, pose extraction can offer a scalable and intelligent method to support data-driven decision-making in modern basketball coaching and player development. This technique plays a critical role in modern sports analytics especially if there are needs to generate insightful evidence about past actions like player behavior, movement efficiency and team dynamics without the need for manual annotation.

International Journal of Artificial Intelligence, vol. 12, no. 2, pp. 101-115, December 2025. **DOI: 10.36079/lamintang.ijai-01202.908**

2.2.1. Python Library and Multi-Metrics for Evaluating Pose Extraction Models

Six basic Python libraries (including os, numpy, OpenCV, pathlib, mediapipe and scipy.stats) were used to implement the objectives of the research described in this paper. The os was used to enable functions and activities or tasks like file and directory management to interact with the operating system. The numpy package was used for efficient computation of operations that involve arrays and mathematical functions. OpenCV (cv2) library was used for the processing, extraction and manipulation of images in the basketball videos. The pathlib was used to handle file system paths and make path manipulations intuitive and readable. The mediapipe is a framework by Google designed for building multimodal models on pose detection involving audio and video footage. The scipy.stats is a sub-package of SciPy and it was used for calculating entropy, mean and standard deviation. Entropy, mean and standard deviation are statistical tools that can assist coaches, players and analysts to understand player movements in sports footage especially when analyzing body joint coordinates or pose vectors across multiple frame samples in the context of pose recognition in basketball videos [29] [30].

Entropy explicitly relates to pose extraction from basketball videos. Entropy identifies moments of pose ambiguity such as transition between dribble and shoot. Such transition can signify specific areas whereby the model is uncertain about the variability and complexity of the body movements of players over a given period of gameplay in basketball videos. The pose (or positions) of players can form patterns representing different actions like dribbling, shooting and defending when pose key points (or joints) are tracked frame-by-frame. Entropy is a statistical metric that measures how unpredictable or diverse such joint positions are within a short video clip.

Mathematically, entropy is expressed as follows [7]:

$$H(P) = -\sum_{i=1}^{n} p_i \log_2(p_i) \tag{1}$$

where,

P = {$p1$, $p2$, ..., pn} is the probability distribution over pose classes (i.e., shoot, pass, defend, dribble, etc.), pi is the predicted probability of the ith pose and n is the total number of pose classes.

High entropy is usually interpreted to indicate a wide range of motion and varied poses. This may correlate to dynamic and complex movements characteristically associated with skilled players or intense gameplay. Conversely, low entropy suggests repetitive and minimal body movements repeatedly occurring in basketball videos. Stationary dribble, free throw preparation, standing in a set defense (e.g., defender holding an opponent in a static posture while watching the basketball) and corner shooter are some common basketball actions that can lead to low pose entropy.

Mean pose vector (μ) measures the average body joint position across multiple players/frames in a given time interval [29]. Mathematically,

$$\mu = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_i \tag{2}$$

where,

X_i is the pose vector for the ith frame (e.g., joint coordinates), N is the number of frame samples and μ is the average pose vector.

This metric is used in basketball videos to identify the average of distinctive pose for a shooting action over many examples. For instance, mean pose vector represents the average position of joints such as elbows, knees and shoulders across multiple poses such as across several shooting actions and dribbling moves in a basketball game play. This metric is used because of its capability to serve as a reference pose for a given action and its capability to create templates and compare different player techniques.

Standard Deviation (σ) measures the variability of pose data (movement of players) [29]. This metric is useful to detect differences and inconsistencies in performance. Mathematically, standard deviation (σ) is expressed as follows:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \mu)^2} \tag{3}$$

where,
X_i is each pose vector, μ is the mean pose vector and N is the number of frame samples. Standard deviation is used to assess the level of consistency of a player's shooting posture over multiple attempts.

In essence, we use this metric to measure and quantify the degree at which individual poses deviate from the mean pose. Low σ (small) signify consistent movement across frames or players and high precision and repeatability in poses. Conversely, high σ (large) indicates high variability or inconsistency in poses. Such value can also correlate to diverse movement styles, errors, and noise in poses.

The combined interpretation of low mean (μ) and low standard deviation (σ) in pose recognition connote that the player's movement is centered around a typical pose and very consistent. This might mean that a professional player perform jump shot with little variation. On the other hand, low mean (μ) and high (σ) imply that there is a central pose in the videos linked to inconsistent players that repetitively performing similar actions but in different ways. Statistically, high standard deviation (σ) and regardless of the value of the mean (μ) may indicate unclear classification, camera noise, pose estimation errors and diverse styles of playing basketball game in the video. The practical usage of the above metrics in basketball analytics can be grouped into four categories. They can be used for skill assessment because lower standard deviation (σ) suggests better form consistency. They can also be used for injury detection. The reason is that sudden increase (or surge) in standard deviation (σ) of pose can suggest impaired movement in some context. Besides, the above metrics can be used to complement training feedback especially in comparing the distribution of player's pose to mean pose of elite athletes. Finally, the above metrics can be used to complement visual diagrams or comparisons of players based on their performance. In this case, high standard deviation (σ) in predictions of pose model has to do with uncertainty and data noise. Analysts may need the coordinates of players to objectively discern and capture how much a player's posture changes within the duration of every game. Such information can be used to advise and discriminate team members on their different playing styles and skill levels required to win highly ranked opponents in competitive basketball games.

2.2.2. Challenges with Pose Extraction from Basketball Videos

There are ethical issues that commonly associate with pose extraction from basketball videos [26]. Using tools like MediaPipe Pose, OpenPose and similar frameworks often come with technical and practical challenges due to the fast-paced, cluttered and dynamic nature of the basketball sport. Occlusion of body parts is a major factor that can constitute technical and practical challenges. For instance, players often block each other during defense and crowding in the paint. This can make the parts of the body completely invisible when viewed behind the other players. Hence, pose models may fail to accurately detect hidden joints and inevitably output inaccurate key points. Basketball games require fast movements which can occasionally cause motion blur. Dunking, sprinting and sudden cuts can inadvertently introduce motion blur in frames. Unfortunately, blur reduces contrast and sharpness of videos and eventually degrading joints detection accuracy. Low resolution, pan, tilt, crowded scenes and zoom can vary across frames in some basketball videos. These features can create confusion, inconsistent scaling and framing such that it could be hard for action recognition models to maintain tracking continuity. Other minor issues such as inconsistent court lighting and glares from polished floors can limit the accurately of action recognition models.

Additionally, most of the existing models are computationally ineffective [27] [6]. Some of them that focus on geometric features also require further customization before they can classify pose or go beyond their initial designs (such as detection of appearance, ball and court context of basketball games). Thus, it is imperative for models designed to analyze, recognize and categorize actions in basketball videos to avoid direct classification but strictly adhere to the creation of privacy-friendly and interpretable features that are reused in downstream tasks involving action recognition within basketball videos. Unfortunately, most existing models are not privacy-friendly analytics, versatile and explainable [20]. Consequently, they leverage on appearance, ball and court context. Thus, this paper proposes a privacy-friendly model for pose extraction across numerous basketballs. The purpose of this technique is to protect all the players, sporting infrastructure and other stakeholders captured in all the input basketball videos. The model also exclude the tracking of physiological and environment attributes to minimize data collection while ensuring transparency and compliance with privacy regulations in sports. The model collects player-related data and excludes personally identifiable information that be used to identify players, infrastructure or contravene good data governance. In essence, pose extraction has continued to generate new and advanced application techniques in the domain of action recognition and in relation to injury prevention analysis, virtual coaching and the development of intelligent basketball simulation systems [25].

3. Methodology

Python programming language and above associated libraries were used to implement the objectives of this research paper. This study used a computational approach to analyze pose-based entropy in basketball videos in other to quantify motion variability and identify patterns that may distinguish player skill levels or gameplay styles. The model encompassed six stages namely loading input basketball videos, basketball video preprocessing, pose estimation, clips aggregation, entropy computation and Output data storage shown in Figure 1 below.

Twenty-two (22) basketball videos were collected and stored in a predefined directory. The model was fine- tuned to recognize a wide range of videos with the following extensions: .mp4, .avi, .mov and

.mkv. Each video was divided into short clips of 1 second duration to ensure consistent temporal segmentation during entropy analysis across all the 22 input basketball videos used to evaluate the model. The frame rate (FPS) of each video was dynamically extracted to determine the number of frames per clip. Thereafter, the MediaPipe Pose solution by Google was used for pose extraction. This technical resource detected 33 body landmarks per frame and provided normalized 2D coordinates (x, y) for each visible joint. The model stored the joints coordinates each time pose joints (or key points) was successfully detected for each frame in a video clip. Conversely, it appended a zero-filled array to maintain clip length consistency whenever it detected no joint.

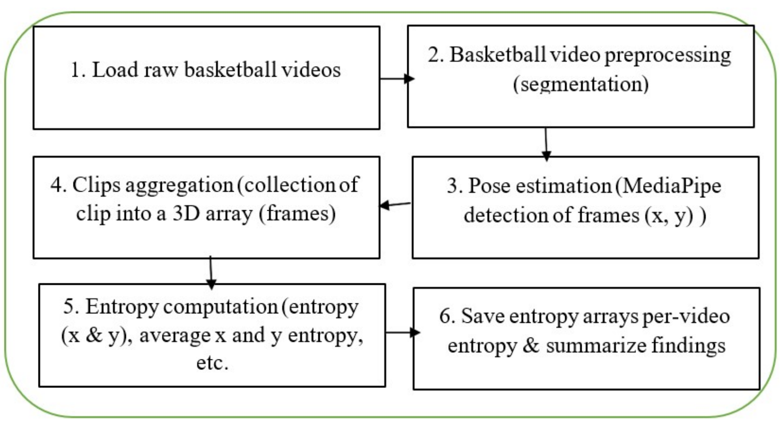

Figure 1. Model Architecture

Each 1 second clip is assumed to constitute several frames (based on the video's FPS) and it was represented by a list of joint arrays. These arrays were aggregated into a 3D tensor of shape (frames × 33 × 2). Entropy computation was performed on them in stepwise manner. The x and y coordinates across all joints and frames were flattened into two 1D arrays. Thereafter, each array was binned into 30 equal-width bins in the normalized range [0, 1] and a histogram density was computed for them. The model import Shannon entropy from the *scipy.stats.entropy* function and use it to separately calculate for both the x and y distributions. The final pose entropy score for the clip was derived by averaging the x and y entropy values. This process reflects the spatial distribution and variability of body movements of players captured in the basketball videos during each clip. It enables the model to automatically perform quantitative analysis of motion complexity in each basketball videos.

The model then saved an array of entropy values (i.e., one per clip) that it computed for each video processed in a designated results directory. These arrays can be used for tasks such as segmenting redundant against informative clips, comparing motion patterns between amateur and professional players and optimizing video review for coaching and analysis on a larger note. Essentially, the above procedures were iterations of automated processes and all the input basketball videos were kept in a specified directory. The script ensured robust handling of these videos with missing frames or undetectable landmarks and logged each step of processing for traceability. Some of the key results obtained from the above procedures are discussed below.

4. Finding and Discussion

Figures 2 to 9 demonstrate some of the key findings from the above model while Tables 1 and 2 display the model's results whereby detected posture transitions were algorithmically scored to indicate the importance and relevance of video segments for detailed analysis. Further analysis reveals that basketball_Video2, basketball_Video3 and basketball_Video9 exhibit significant transitions and elevated priority scores to suggest that their segments merit further examination.

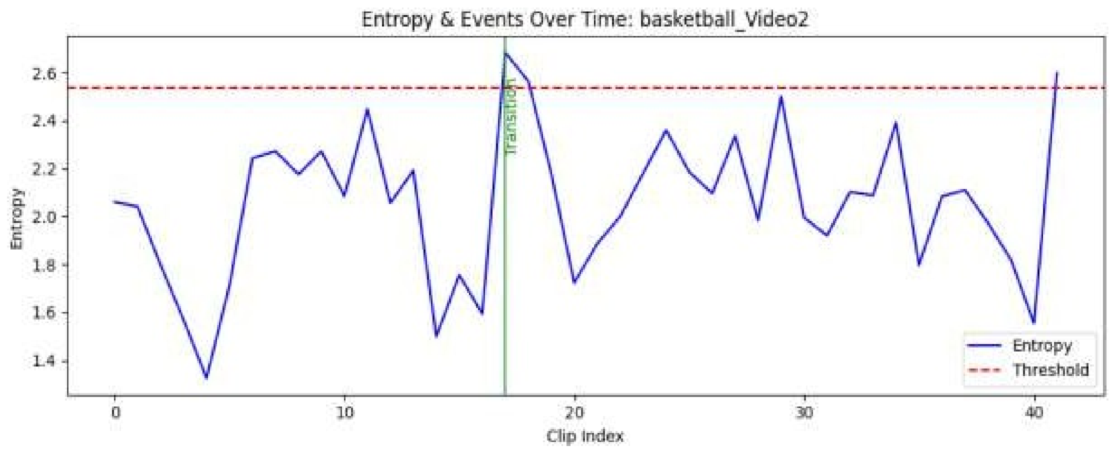

Figure 2. Sharon Entropy for Basketball_Video2

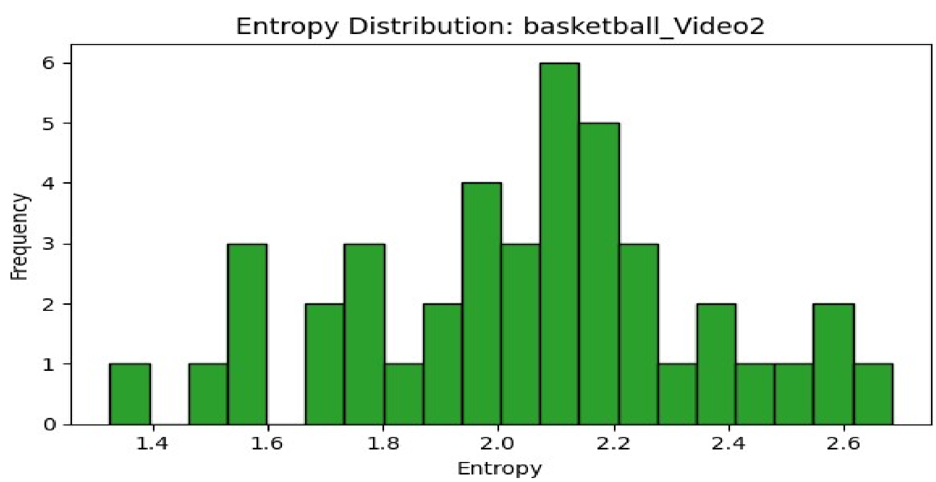

Figure 3. Entropy histogram for Basketball_Video2

From Figure 2 to Figure 9, the transition indices for basketball_Video2 occur at clip 17 while its entropy statistics showing a mean of 2.05 and a standard deviation of 0.31. For basketball_Video3, the transition occurs earlier at clip 1 with a mean entropy of 1.94 and a standard deviation of 0.20. Meanwhile, basketball_Video9 has transitions at clips 7 and 11. Its entropy statistics indicate a mean of 1.96 and a standard deviation of 0.33.

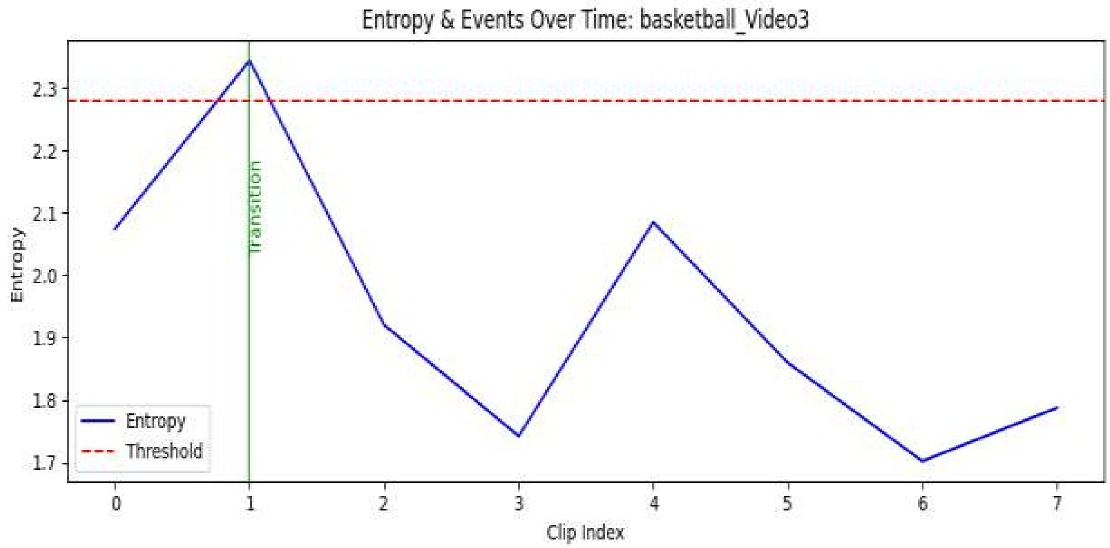

Figure 4: Entropy for Basketball_Video3

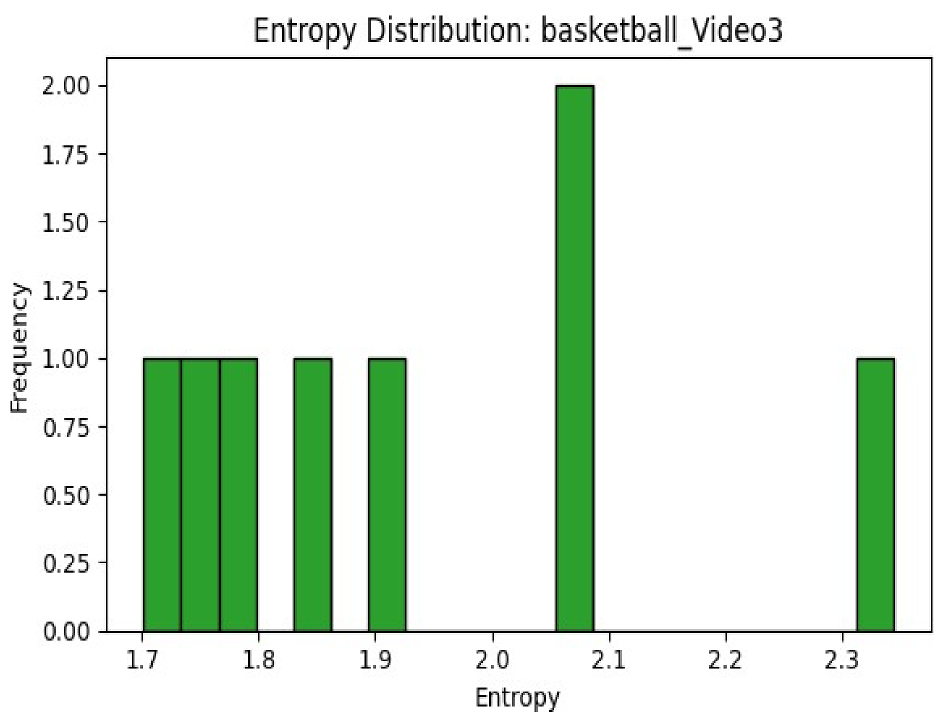

Figure 5: Entropy histogram for Basketball_Video3

International Journal of Artificial Intelligence, vol. 12, no. 2, pp. 101-115, December 2025. **DOI: 10.36079/lamintang.ijai-01202.908**

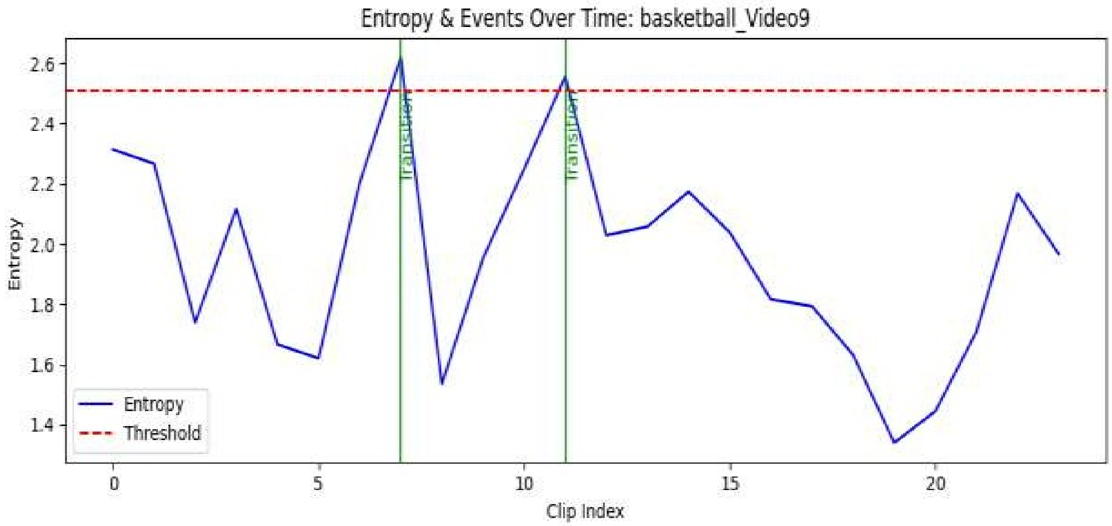

Figure 6: Entropy for Basketball_Video9

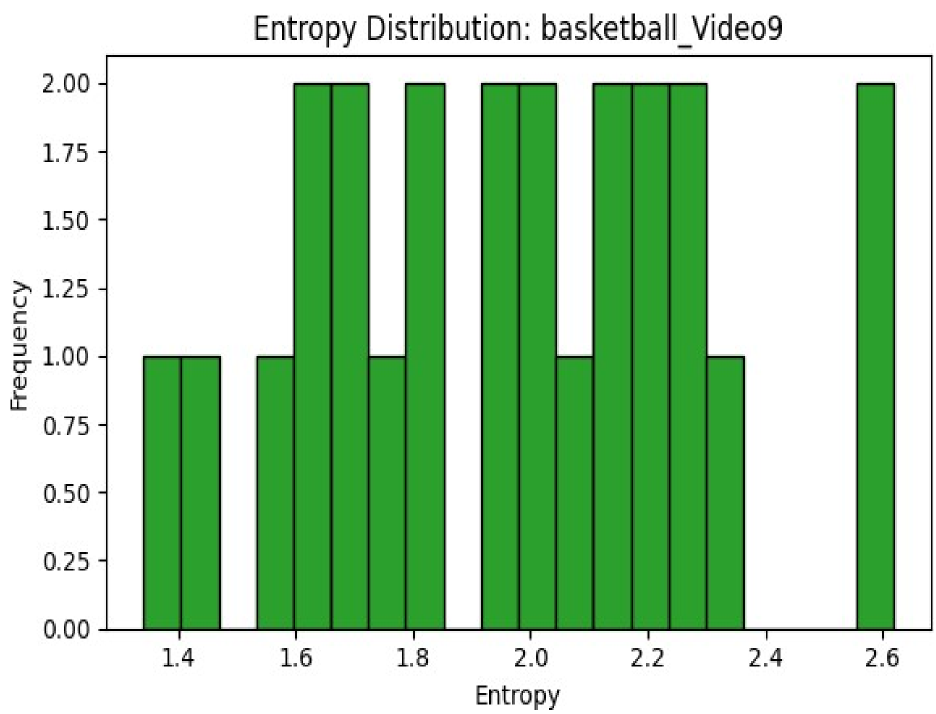

Figure 7: Entropy histogram for Basketball_Video9

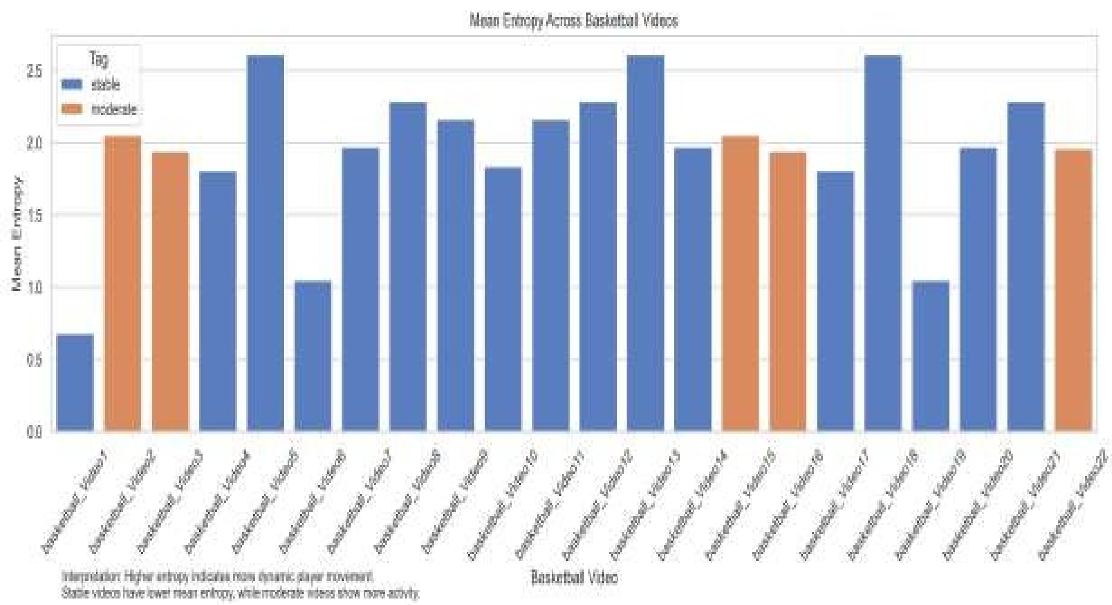

Figure 8. Mean Entropy for all Basketball Videos

Table 1. Summary of Entropy Across 22 Basketball Videos

Video	Tag	Transition Type	Transitions	Review Priority	Mean Entropy	Std Entropy	Skew Entropy
basketball_Video10	stable	no_transition	0	Low	0.679206	0.483584	-0.4872
basketball_Video11	stable	no_transition	0	Low	2.16102	0.57607	-1.82659
basketball_Video12	stable	no_transition	0	Low	1.834639	0.537149	-1.74205
basketball_Video13	stable	no_transition	0	Low	2.16102	0.57607	-1.82659
basketball_Video14	stable	no_transition	0	Low	2.285289	0.478649	-0.31912
basketball_Video15	stable	no_transition	0	Low	2.608631	0.267977	-0.93303
basketball_Video16	stable	no_transition	0	Low	1.967744	0.652832	-1.64622
basketball_Video17	stable	no_transition	0	Low	1.834639	0.537149	-1.74205
basketball_Video18	stable	no_transition	0	Low	2.16102	0.57607	-1.82659
basketball_Video19	stable	no_transition	0	Low	1.834639	0.537149	-1.74205
basketball_Video1	stable	no_transition	0	Low	1.834639	0.537149	-1.74205
basketball_Video20	stable	no_transition	0	Low	1.967744	0.652832	-1.64622
basketball_Video21	stable	no_transition	0	Low	2.608631	0.267977	-0.93303
basketball_Video22	stable	no_transition	0	Low	2.285289	0.478649	-0.31912
basketball_Video2	moderate	middle	1	Medium	2.052863	0.305494	-0.17746
basketball_Video3	moderate	early	1	Low	1.939146	0.202709	0.703646
basketball_Video4	stable	no_transition	0	Low	1.804156	0.217476	1.079796
basketball_Video5	stable	no_transition	0	Low	2.608631	0.267977	-0.93303
basketball_Video6	stable	no_transition	0	Low	1.047717	0.256401	-0.20943
basketball_Video7	stable	no_transition	0	Low	1.967744	0.652832	-1.64622
basketball_Video8	stable	no_transition	0	Low	2.285289	0.478649	-0.31912
basketball_Video9	moderate	moderate	2	Medium	1.958646	0.327994	0.069801

International Journal of Artificial Intelligence, vol. 12, no. 2, pp. 101-115, December 2025. **DOI: 10.36079/lamintang.ijai-01202.908**

From Table 1, the mean entropy and standard deviation (Std) for basketball_Video21, basketball_Video5 and basketball_Video15 are 2.6086 and 0.268 respectively. A higher mean entropy generally indicates more dynamic gameplay that is often characterized by frequent player movements or positional changes. Conversely, a lower standard deviation reflects consistent behavior (e.g., a steady pace with minimal fluctuations in intensity). Interestingly, these three videos were tagged as 'stable', assigned a 'low' review priority and exhibited 'no transitions' despite their high entropy values. Similarly, the mean entropy is high (i.e., 1.958646) and standard deviation (entropy) is low (i.e., 0.327994) during pose recognition in basketball_Video9. These combined results suggest that pose classes are too similar with many overlaps in this video. The low standard deviation of entropy also indicates that the level of uncertainty of the model is consistent across all frames and poses in the video. Technically, there is need for the coach or analysts to critically examine camera angles, occlusions (e.g., motion blur) and similar poses (e.g., “prepare to shoot” and “shoot” and merge them to form one pose) in other to remove ambiguity or achieve clearer pose boundaries. These observations suggest that a high mean entropy does not inherently signal critical gameplay moments or the need for closer review. That is not all high- motion sequences require detailed analysis by coaches, analysts and players especially if the intensity of the gameplay remains consistent or uniform throughout the events.

Figure 9 illustrates the scatter diagram that visually compare the transitions and priority score of the input datasets. The results prove that entropy-based system can flag transitions in player posture. The X-axis refers to the number of pose transitions detected in basketball videos using entropy arrays and statistical features (mean and standard deviation). The Y-axis indicates priority score (1=Low, 3=High) that was algorithmically assigned by the model during analysis. These two scores represent the importance, intensity and relevance of a segment in the video based on pose changes.

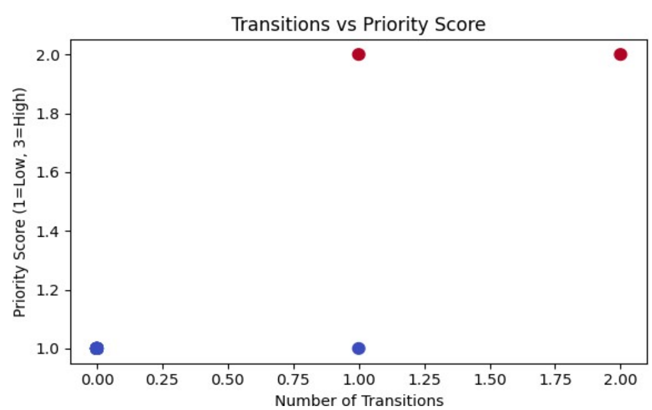

Figure 9. Transitions versus Priority Score for all Basketball Videos

There are 5 points plotted in Figure 9 with two different colors. Blue dots signify priority score of 1 (low priority) while red dots indicate priority score of 2 (medium priority). The above results indicate that higher transitions can indicate higher priority. The red dots (Priority Score = 2) appear at 1 and 2 transitions. These suggest that the scene becomes more dynamic and potentially more important as the number of pose transitions increases. Hence, it has a higher priority score. The results also mean that some transitions do not have high priority. The presence of blue at 1 transition indicates that transition count alone is not enough to determine priority of pose in basketball games. Zero transitions usually have low priority. Consistent blue at 0 transitions implies that low pose activity (static posture) is generally less important.

Table 2 expands the insights from Figure 9 to further show how the metrics help validate the scoring logic and reveal a consistent correlation between motion and assigned priority. Essentially, our analytical model has successfully distilled the review process from 22 basketball videos down to just three critical ones that truly warrant closer examination. This targeted approach does not only enhance efficiency but also significantly conserves valuable time and resources by filtering out 19 basketball videos that are likely redundant or of lower priority. Such precision in identifying key moments underscores the effectiveness of our method in reducing workload, prioritizing impactful gameplay events and ultimately enabling coaches, players and sport analysts to have more focused and resource-savvy review workflow.

Table 2. Analysis of Pose Transitions in Basketball Videos

Transitions	Priority Score	Color	Interpretation
0	1	Blue	No transitions, low-priority segment
1	1	Blue	One transition, still low-priority
1	2	Red	One transition, higher relevance
2	2	Red	Two transitions, medium-priority
0	1	Blue	Repetition of low-priority with no transitions

5. Conclusion

Models on automatic recognition of actions in basketball videos must be privacy-friendly analytics, versatile and explainable so that coaches, players and analysts will not waste resources needed to improve performance on trivial review. Unfortunately, most of the existing models depend on appearance, court visuals and sensor data and they frequently raise potential privacy concerns. Some of them that adopt Deep Learning approaches like 3D CNNs and IMU classifiers are computationally heavy and often lack intuitive interpretability compared to entropy metrics (mean and standard deviation) that can offer clearer insights into action recognition. There are significant numbers of them typically trained on limited pose/action datasets under controlled conditions. So, they often fail to capture the full variability of basketball gameplay. Additionally, they generally treat all clips equally without prioritization, making them ineffective at filtering out redundant videos and hindering analysts from focusing their review efforts efficiently.

Thus, this study presents an entropy-based analytical model for automatic recognition of key actions in basketball videos to optimize the video review process. By using Python programming language to implement and analyze entropy arrays, mean and standard deviation derived from 22 basketball game videos, the model effectively narrows down these videos into just three critical videos (with mean=1.96 & std=0.33, mean=2.05 $ std=0.31, mean=1.94 & std=0.20) requiring detailed examination. This targeted reduction enhances efficiency by significantly conserving time and resources, eliminating 19 videos deemed redundant and of lower priority. The above approach has demonstrated high precision in identifying impactful gameplay moments and addresses a long-standing challenge in basketball analytics and reduction of workload while maintaining review accuracy. This method has also created privacy-friendly and offered coaches, players and sports analysts a focused and resource-efficient framework for performance evaluation and strategic planning in basketball sports.

Furthermore, like every model on action recognition in basketball videos, our model would obviously have some limitations that can be explored to enhance its novelty in future time. For instance, entropy alone may oversimplify complex motion patterns within some basketball videos. Entropy may not perfectly model temporal dependencies, joint interactions and multi- person dynamics in all categories of basketball videos. The metrics may assume all pose landmarks are equally important but such assumption may not necessarily hold for actions like shooting and dunk. The transition points for basketball_Video2 were identified at clip 17 and its corresponding entropy statistics showing a mean of 2.05 and a standard deviation of 0.31. Transitions occurred early at clip 1 in basketball_Video3 with a mean entropy of 1.94 and a lower standard deviation of 0.20. Basketball_Video9 exhibited transitions at clips 7 and 11. Its entropy values reflecting a mean of 1.96 and a higher variability indicated by a standard deviation of 0.33. These observations suggest that while entropy metrics capture movement dynamics but it is possible that an entropy spike may not absolutely predict the review priority (such as medium, low and high) required to segment basketball videos. This highlights the need for integrating entropy analysis with contextual understanding (like manual review of the prioritized videos) to accurately assess the significance of gameplay events.

International Journal of Artificial Intelligence, vol. 12, no. 2, pp. 101-115, December 2025. **DOI: 10.36079/lamintang.ijai-01202.908**

Thus, our future research would incorporate multivariate entropy, multimodal fusion and mutual information between joints to capture coordinated movements of players in basketball videos. It is also plausible that we use additional modalities that can combine pose entropy with audio, ball trajectory and court position data for richer representation of actions in above context. We intend to pursue the above possibilities in our future research to extend the above findings.

Reference

[1] J. Lu, "Basketball motion recognition model analysis based on perspective invariant geometric features in skeleton data extraction," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 11, 2023.

[2] Q. Zhang, "Lightweight classification method of human action in long video of basketball competition," in *Proc. 8th Int. Conf. Inf. Syst. Eng. (ICISE)*, Dalian, China, pp. 13–21, 2023.

[3] J. Xiao, W. Tian, and L. Ding, "Basketball action recognition method of deep neural network based on dynamic residual attention mechanism," *Information*, vol. 14, no. 1, 2022.

[4] J. Wang, L. Zuo, and C. Cordente Martínez, "Basketball technique action recognition using 3D convolutional neural networks," *Sci. Rep.*, vol. 13, no. 13156, pp. 1–14, 2024.

[5] J. Li and D. Gu, "Research on basketball players' action recognition based on interactive system and machine learning," *J. Intell. Fuzzy Syst.*, vol. 40, no. 2, pp. 2029–2039, 2021.

[6] J. Liu, X. Liu, M. Qu, and T. Lyu, "EITNet: An IoT-enhanced framework for real-time basketball action recognition," *Alexandria Eng. J.*, vol. 110, pp. 567–578, 2025.

[7] L. Zhao and W. Chen, "Detection and recognition of human body posture in motion based on sensor technology," *IEEJ Trans. Electr. Electron. Eng.*, vol. 15, no. 5, pp. 766–770, 2020.

[8] C.-Y. Chen, J.-C. Wang, J.-F. Wang, and Y.-H. Hu, "Motion entropy feature and its applications to event-based segmentation of sports video," *EURASIP J. Adv. Signal Process.*, vol. 2008, Article 460913, 2008.

[9] M. Takahashi, M. Naemura, M. Fujii, and J. J. Little, "Recognition of action in broadcast basketball videos on the basis of global and local pairwise representation," in *Proc. IEEE Int. Symp. Multimedia*, Anaheim, CA, USA, pp. 147–154, 2013.

[10] B. Acikmese and S. Bhatia, "Sensor-driven real-time recognition of basketball goal states using IMU and deep learning," *Sensors*, vol. 25, no. 12, 2020.

[11] Z. Zhang, W. Liu, Y. Zheng, L. Du, and L. Sun, "Learning spatio-temporal context for basketball action pose estimation with a multi-stream network," *Sci. Rep.*, vol. 15, no. 1, 2025.

[12] L. Wu *et al.*, "Ontology-based global and collective motion patterns for event classification in basketball videos," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 7, pp. 2178–2190, 2019.

[13] Y. Bo, "A reinforcement learning-based basketball player activity recognition method using multisensors," *Mobile Inf. Syst.*, vol. 2022, 2022.

[14] X. Ma, Y. Zhang, and Y. Li, "Sensor-driven real-time recognition of basketball goal states using IMU and deep learning," *Sensors*, vol. 25, no. 12, 2020.

[15] L. Wu *et al.*, "Fusing motion patterns and key visual information for semantic event recognition in basketball videos," *Neurocomputing*, vol. 413, pp. 217–229, 2020.

[16] A. Hoelzemann, J. L. Romero, M. Bock, K. Van Laerhoven, and Q. Lv, "Hang-Time HAR: A benchmark dataset for basketball activity recognition using wrist-worn inertial," *Sensors*, vol. 23, no. 13, 2023.

[17] X. Guo, E. Brown, P. P. K. Chan, R. H. M. Chan, and R. T. H. Cheung, "Skill level classification in basketball free-throws using a single inertial sensor," *Appl. Sci.*, vol. 13, no. 9, 2023.

[18] X. Gu, X. Xue, and F. Wang, "Fine-grained action recognition on a novel basketball dataset," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Barcelona, Spain, pp. 2563–2567, 2020.

[19] X. Wu and Y. Xue, "Recognition algorithm of basketball dynamic movement behavior based on multimedia network technology," in *Proc. Int. Conf. Intell. Comput. Mach. Learn. (2ICML)*, Qingdao, China, pp. 51–56, 2023.

[20] S. B. Khobdeh, M. R. Yamaghani, and S. K. Sareshkeh, "Basketball action recognition based on the combination of YOLO and a deep fuzzy LSTM network," *J. Supercomput.*, vol. 80, no. 3, pp. 3528–3553, 2023.

[21] T. Leinster, *Entropy and Diversity: The Axiomatic Approach*. India: Cambridge Univ. Press, 2021.

[22] J. Xie, Y. Chen, and L. Mei, "Research on basketball video action segmentation method based on deep learning," in *Proc. Int. Conf. Comput. Vis., Image Deep Learn. (CVIDL)*, Zhuhai, China, pp. 1372–1376, 2024.

[23] Z. Su, "Designing a basketball action recognition system based on the improved OpenPose algorithm," in *Proc. Asia-Pacific Conf. Image Process., Electron. Comput. (IPEC)*, Dalian, China, pp. 25–29, 2024.

[24] S. Fang, G. Wang, Y. Li, Y. Yu, and J. Li, "Generating deeply-engineered technical features for basketball video understanding," *IEEE Access*, vol. 13, pp. 20667–20677, 2025.

[25] H. B. Wicaksono and Wahyono, "Spatio-temporal action verification of basketball travelling dribble violation using Mediapipe-YOLO-LSTM framework," in *Proc. Int. Workshop Intell. Syst. (IWIS)*, Ulsan, South Korea, pp. 1–6, 2023.

[26] H. Meng, "Basketball training system based on 3D motion intelligent recognition technology," in *Proc. Int. Conf. 3D Immersion, Interact. Multi-sensory Experiences (ICDIIME)*, Madrid, Spain, pp. 384–388, 2023.

[27] X. Cheng and Z. Li, "Research on basketball shooting action recognition and optimization system based on deep learning," in *Proc. Int. Conf. Mechatronics, IoT Ind. Informatics (ICMIII)*, Melbourne, Australia, pp. 546–551, 2024.

[28] R. Ji, "Research on basketball shooting action based on image feature extraction and machine learning," *IEEE Access*, vol. 8, pp. 138743–138751, 2020.

[29] L. Jing, K. Kitahama, M. Yamata, and Y. Hoshino, "Basketball foul model and judgment system proposal," in *Proc. IEEE Global Conf. Consum. Electron. (GCCE)*, Kyoto, Japan, pp. 988–991, 2021.

[30] S. E. Umbaugh, *Digital Image Processing and Analysis: Computer Vision and Image Analysis*. Boca Raton, FL, USA: CRC Press, 2023.